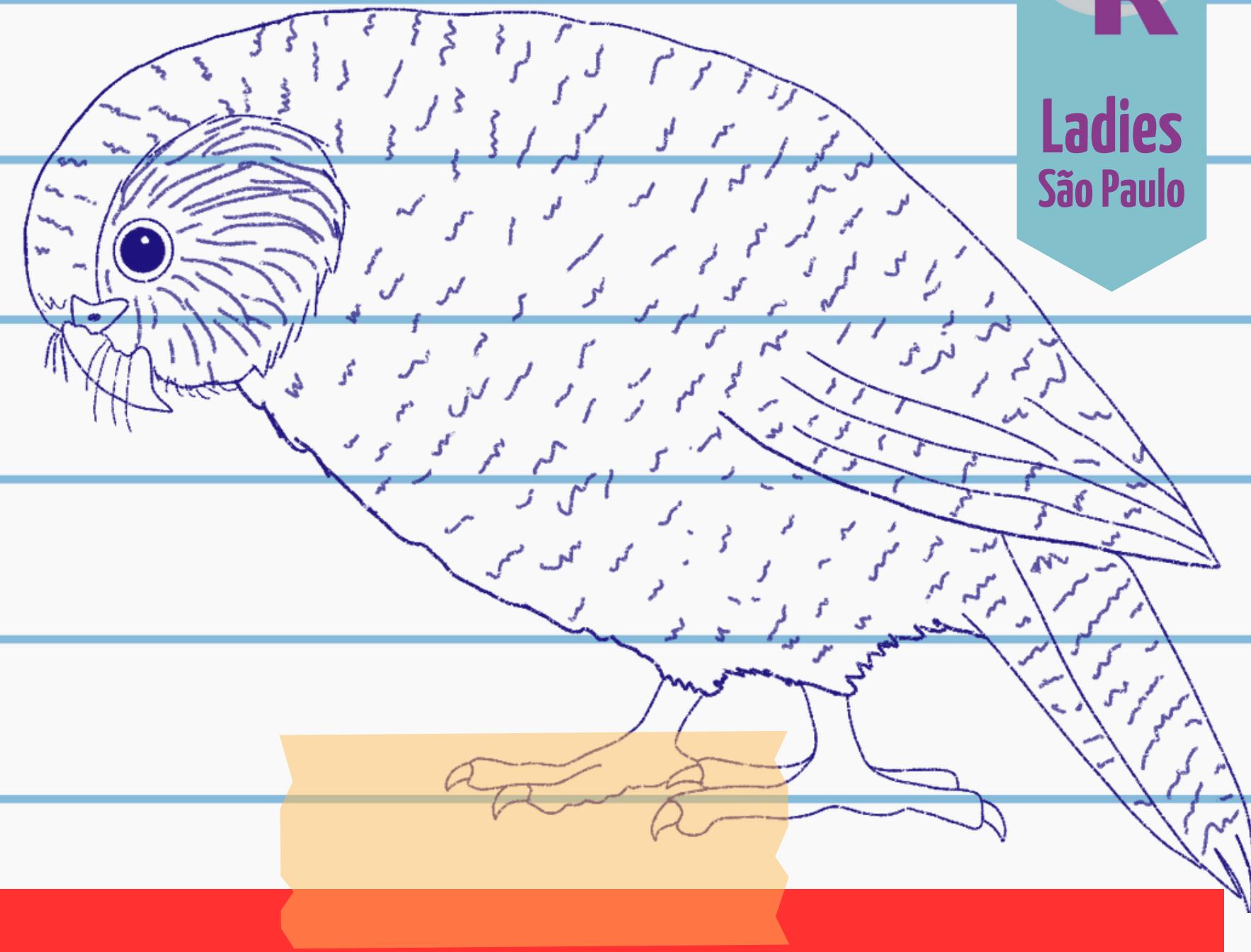
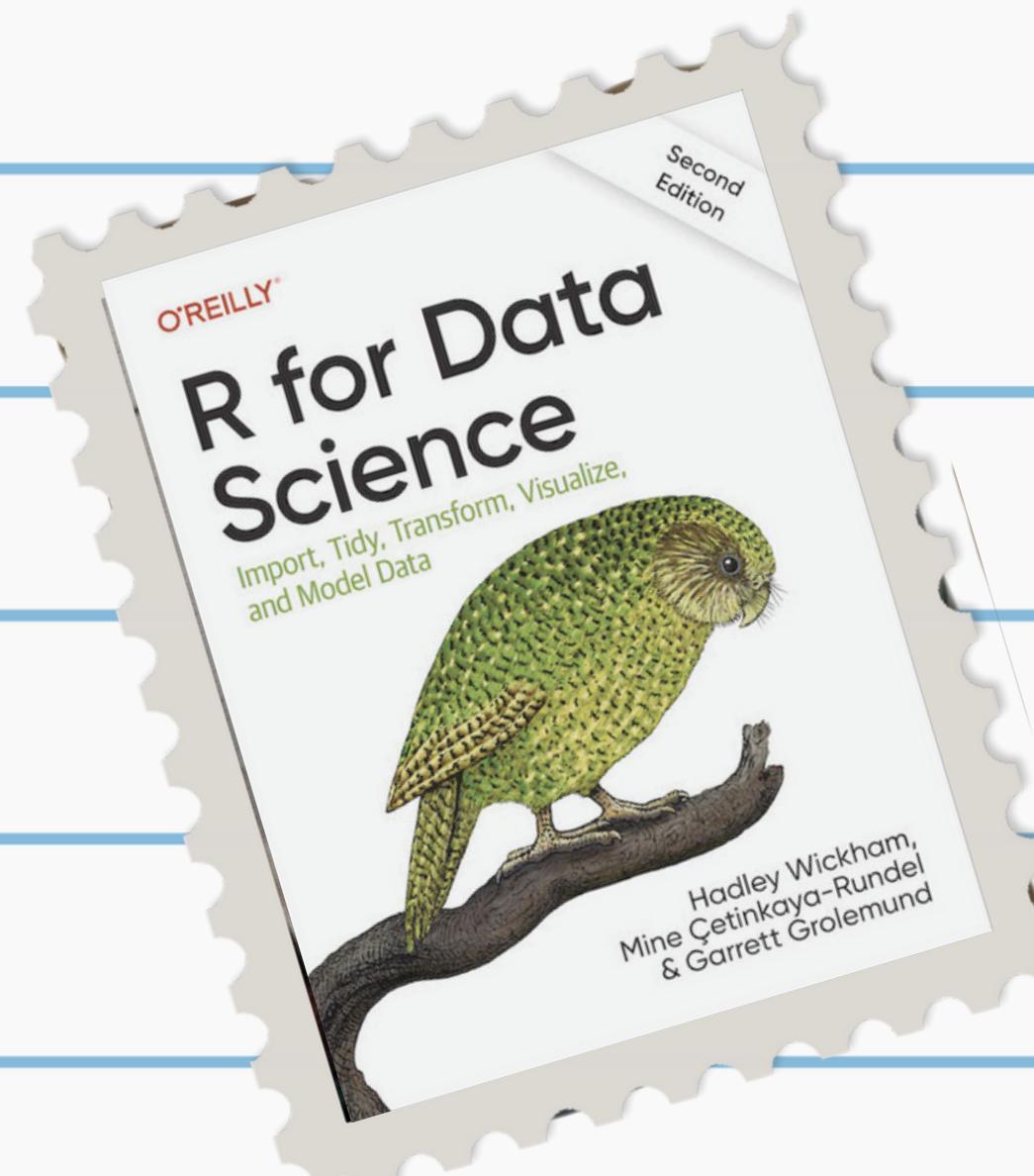


Clube do Livro: R FOR DATA SCIENCE 2ed

Versão traduzida para Português



Encontro 2: Visualização de dados

Dia 02/04, 19h, online

Boas-vindas!

Estrutura do livro

1. Visão geral

2. Visualizar

3. Transformar

4. Importar

5. Programar

6. Comunicar

Fevereiro - Introdução

Março - Visualização de dados

Abri - Fluxo de Trabalho: básico / Scripts e Projetos

Maio - Transformação de dados

Junho - Organização de dados

Julho - Importação de dados

Agosto - Fluxo de Trabalho: estilo de código / buscando ajuda





Recados



<input checked="" type="checkbox"/> Boas-vindas
<input checked="" type="checkbox"/> Prefácio da segunda edição
<input checked="" type="checkbox"/> Introdução
<input checked="" type="checkbox"/> Visão geral
1 <input checked="" type="checkbox"/> Visualização de dados
2 <input checked="" type="checkbox"/> Fluxo de Trabalho: básico
3 Data transformation
4 <input checked="" type="checkbox"/> Fluxo de trabalho: estilo de código
5 <input checked="" type="checkbox"/> Organizando os dados (<i>data tidying</i>)
6 <input checked="" type="checkbox"/> Fluxo de trabalho: scripts e projetos
7 Data import
8 Workflow: getting help
<input checked="" type="checkbox"/> Visualizar
9 <input checked="" type="checkbox"/> Camadas
10 Exploratory data analysis
11 Communication
<input checked="" type="checkbox"/> Transformar
12 <input checked="" type="checkbox"/> Vetores lógicos
13 <input checked="" type="checkbox"/> Números
14 Strings

Atualização sobre a tradução

- Novos capítulos traduzidos
- Marcamos os capítulos já traduzidos com 
- Recebemos sugestões de correções, obrigada:
 - Ludimila Juliele
 - Karina Ferrara Barros
 - Michele Caramigo



Boas vindas ao fórum do Clube do Livro!

annoncios · beatrizmilz

is:open

x

Sort by: Latest activity ▾

Label ▾

Filter: Open ▾

New discussion

Categories



Discussions

View all discussions

↑ 1



Leitura - Capítulo 'Visualização de dados'

Leitura



0

Anúncios

Dúvidas

Exercícios do livro

Fui lá e fiz!

Geral

Votação

↑ 1



Exercícios - Capítulo 'Visualização de dados'

Exercícios



0

↑ 1



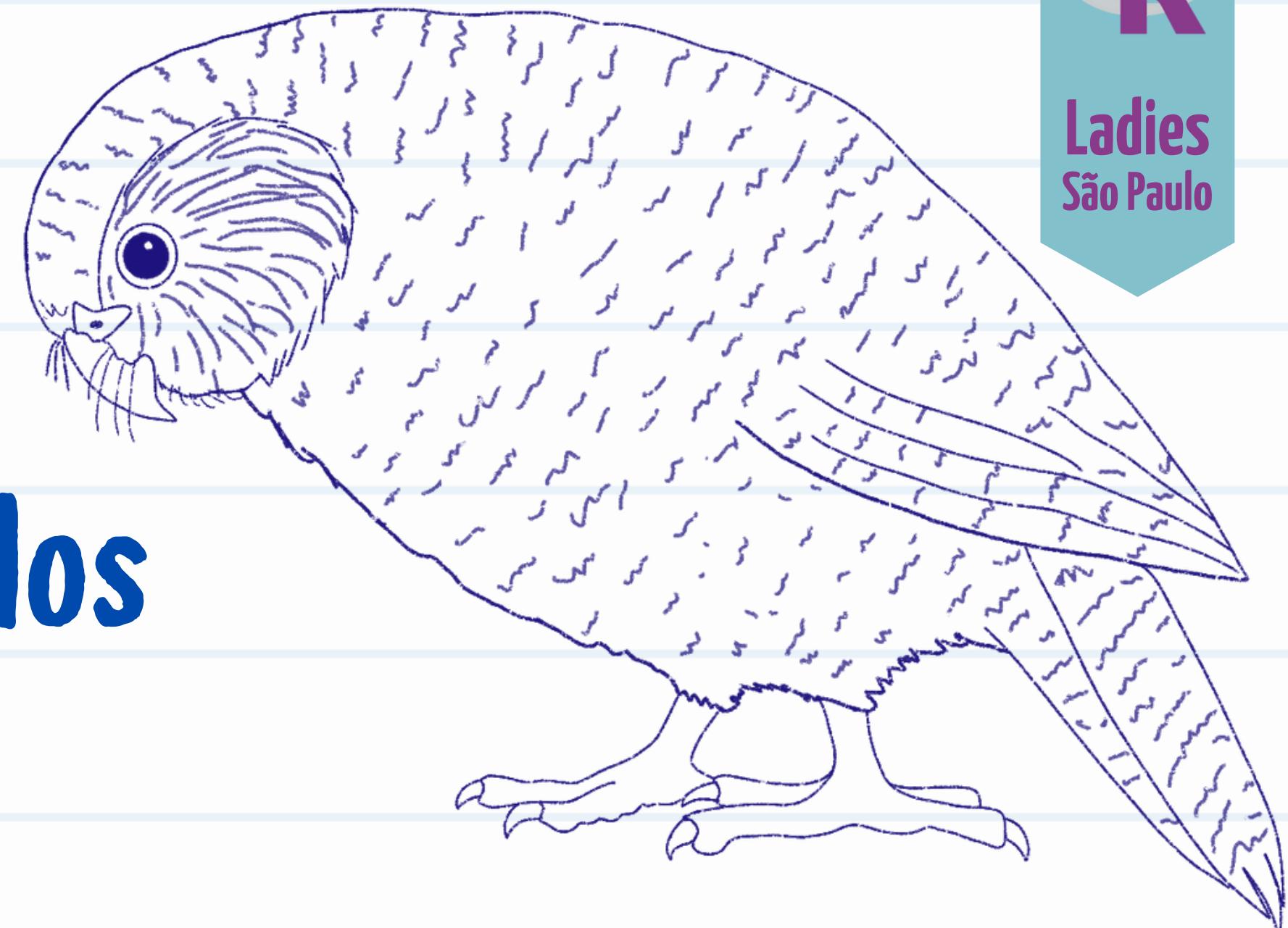
Boas vindas ao fórum do Clube do Livro!

beatrizmilz announced on Jan 28 in Anúncios



0

<https://github.com/R-Ladies-Sao-Paulo/2024-clube-do-livro-r4ds/discussions>



1. Visualização de Dados

Agradecimentos:

Tradução: Bruno Mioto

Revisão: Marcus Ramalho e Beatriz Milz

Instalar o R e o RStudio



Ou use a versão na nuvem:

<https://posit.cloud>

Pacotes

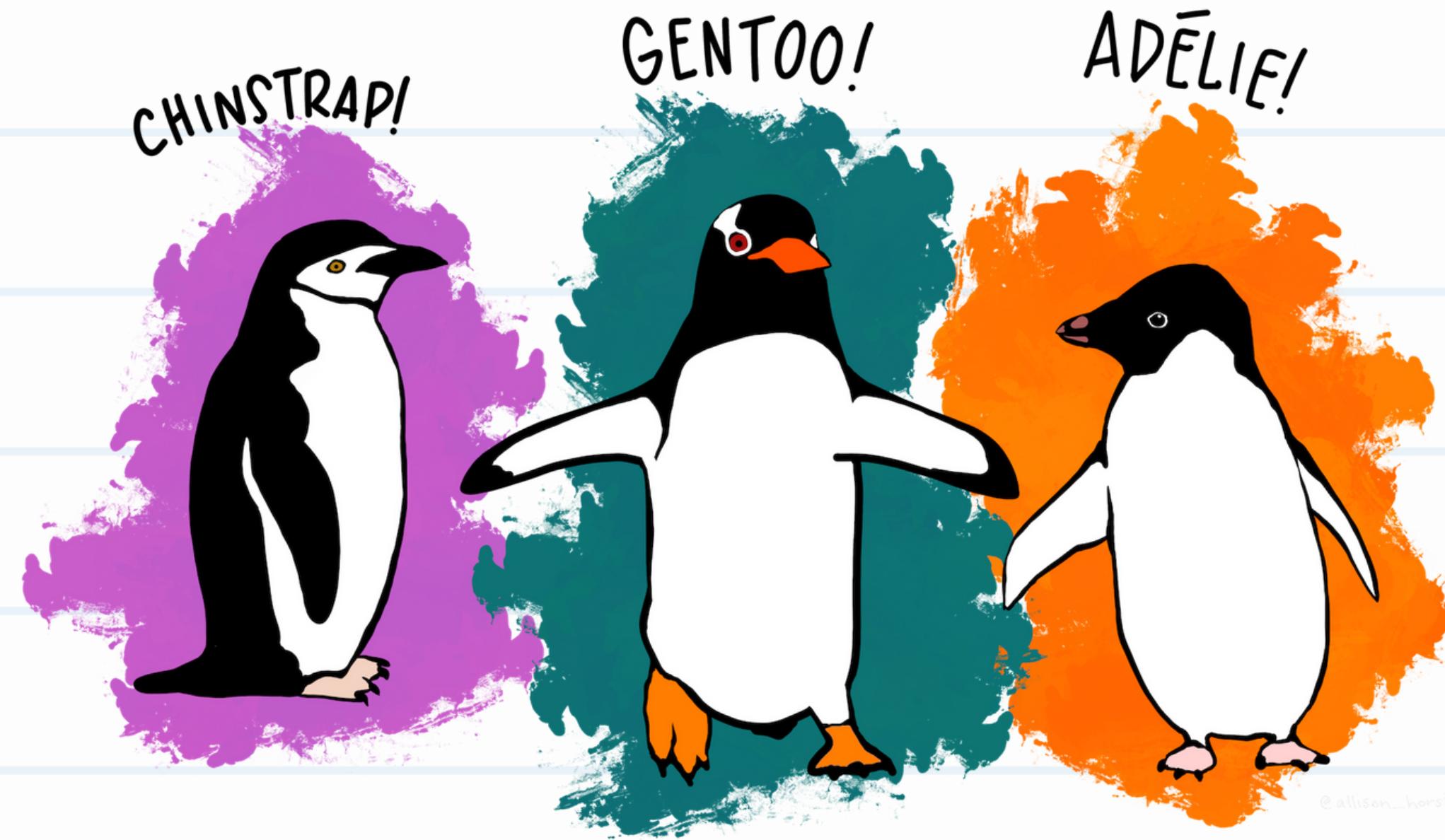
```
install.packages("tidyverse")
install.packages("dados")
install.packages("ggplot")
```

Pacote “Dados”

Datasets

- casas
- premios_gerentes
- pessoas
- salarios
- dados_iris
- mtcarros
- dados_starwars
- veiculos
- diamante
- milhas
- voos
- **pinguins**
- etc...

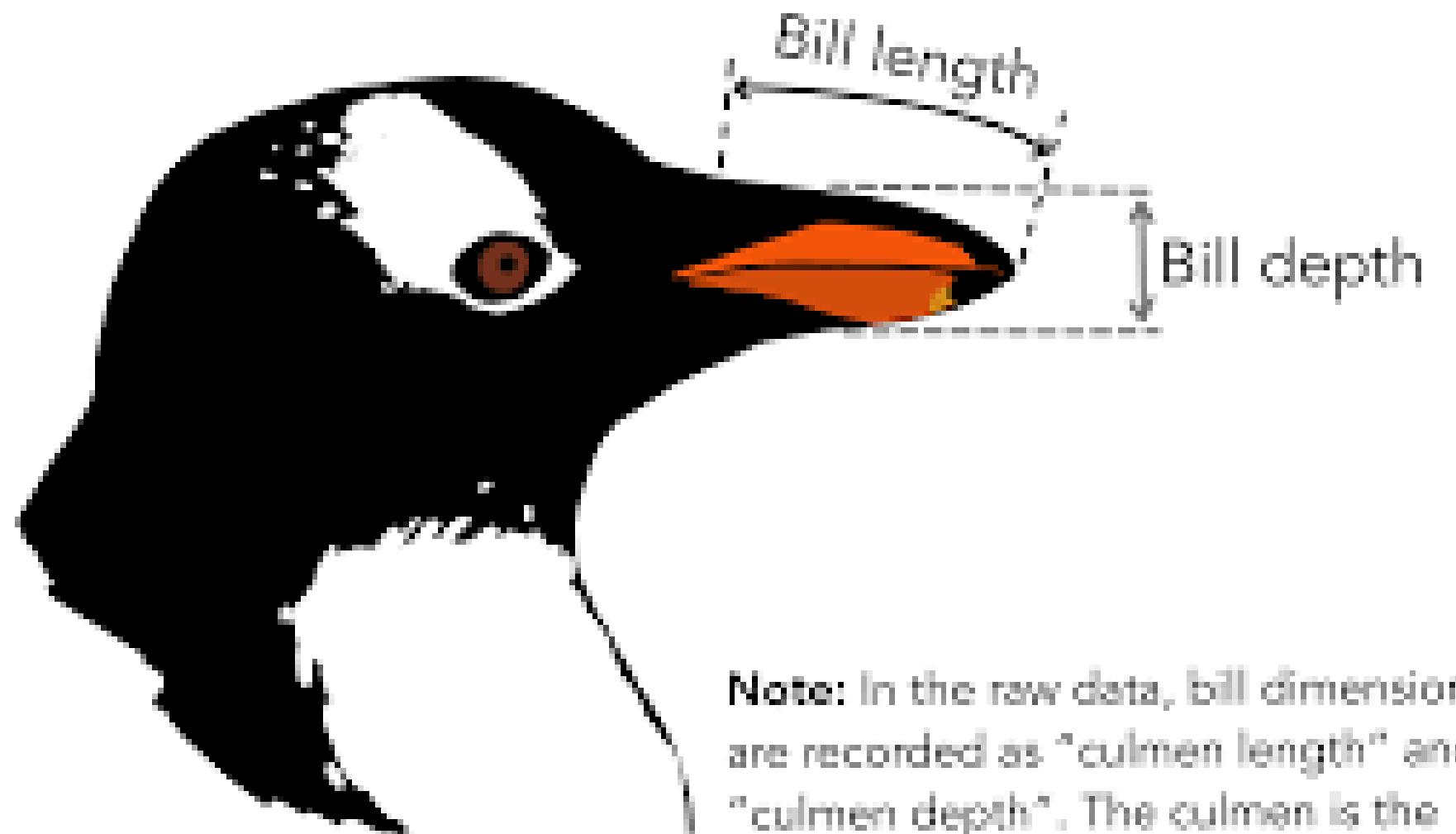
Pinguins



Pinguins

dados::pinguins

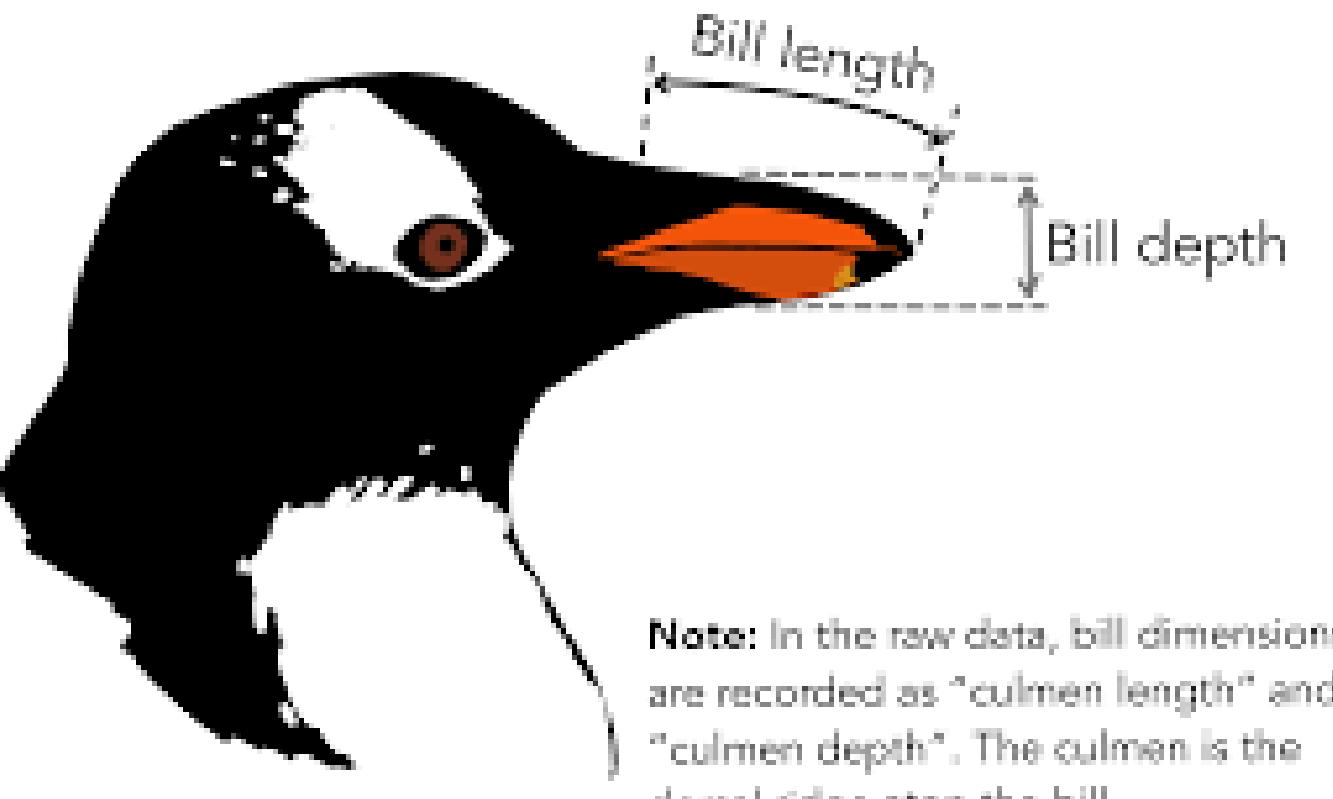
?pinguins



Note: In the raw data, bill dimensions are recorded as "culmen length" and "culmen depth". The culmen is the dorsal ridge atop the bill.

Pinguins

glimpse(pinguins)



Variáveis

1. Variáveis Quantitativas (numéricas)

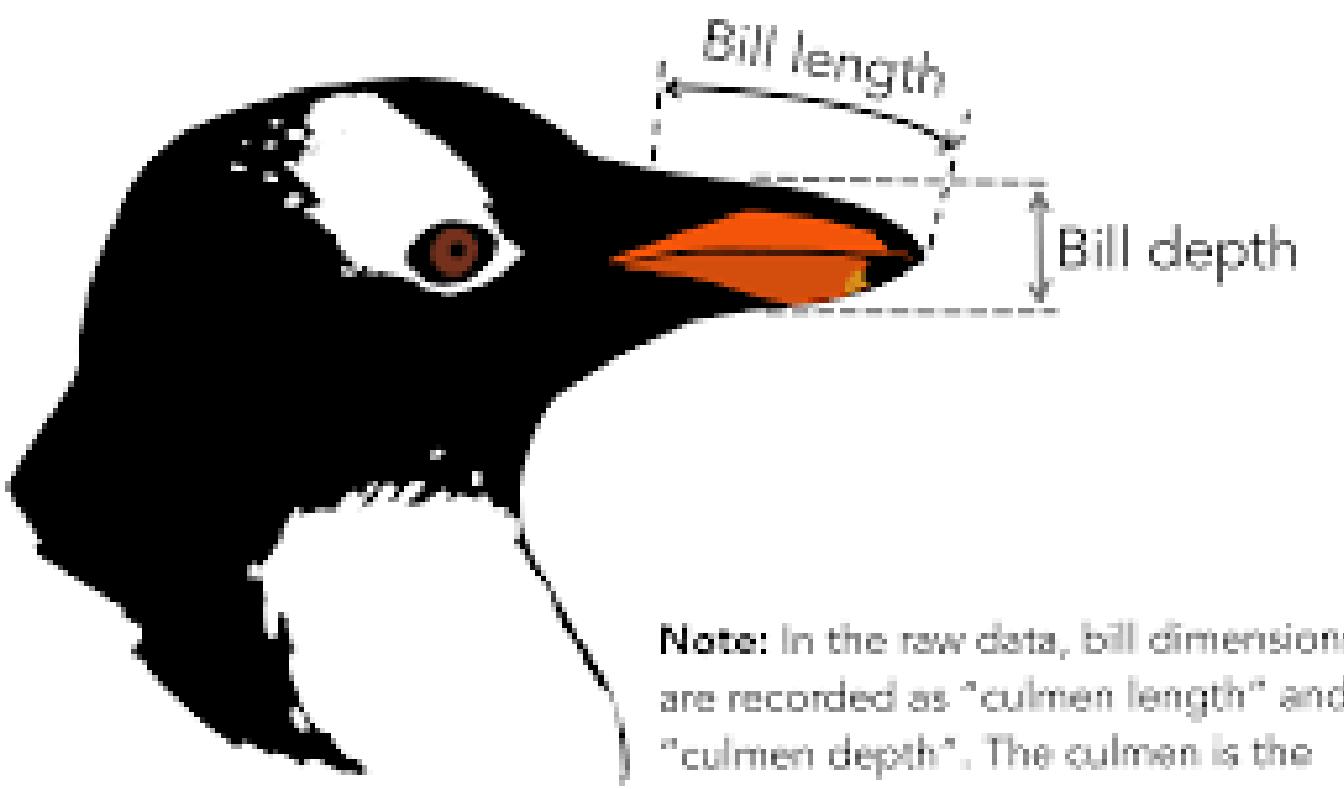
Discretas

Contínuas

2. Variáveis Qualitativas (categóricas)

Nominais

Ordinais



Pinguins

Variáveis

\$ especie

\$ ilha

\$ comprimento_bico

\$ profundidade_bico

\$ comprimento_nadadeira

\$ massa_corporal

\$ sexo

\$ ano

Variáveis

Classifique as variáveis abaixo:

- \$ especie _____ Pinguim-de-adélia, Pinguim-de-adélia, Pinguim-de...
- \$ ilha _____ Torgersen, Torgersen, Torgersen, Torgersen, Torg...
- \$ comprimento_bico _____ 39.1, 39.5, 40.3, NA, 36.7, 39.3, 38.9, 39.2, 34...
- \$ profundidade_bico _____ 18.7, 17.4, 18.0, NA, 19.3, 20.6, 17.8, 19.6, 18...
- \$ comprimento_nadadeira _____ 181, 186, 195, NA, 193, 190, 181, 195, 193, 190,...
- \$ massa_corporal _____ 3750, 3800, 3250, NA, 3450, 3650, 3625, 4675, 34...
- \$ sexo _____ macho, fêmea, fêmea, NA, fêmea, macho, fêmea, ma...
- \$ ano _____ 2007, 2007, 2007, 2007, 2007, 2007, 2007, 2007, ...

Variáveis

Classifique as variáveis abaixo:

\$ especie

Categórica N Pinguim-de-adélia, Pinguim-de-adélia, Pinguim-de...

\$ ilha

Categórica N Torgersen, Torgersen, Torgersen, Torgersen, Torg...

\$ comprimento_bico

N. Contínua 39.1, 39.5, 40.3, NA, 36.7, 39.3, 38.9, 39.2, 34...

\$ profundidade_bico

N. Contínua 18.7, 17.4, 18.0, NA, 19.3, 20.6, 17.8, 19.6, 18...

\$ comprimento_nadadeira

N. Discreta 181, 186, 195, NA, 193, 190, 181, 195, 193, 190, ...

\$ massa_corporal

N. Discreta 3750, 3800, 3250, NA, 3450, 3650, 3625, 4675, 34...

\$ sexo

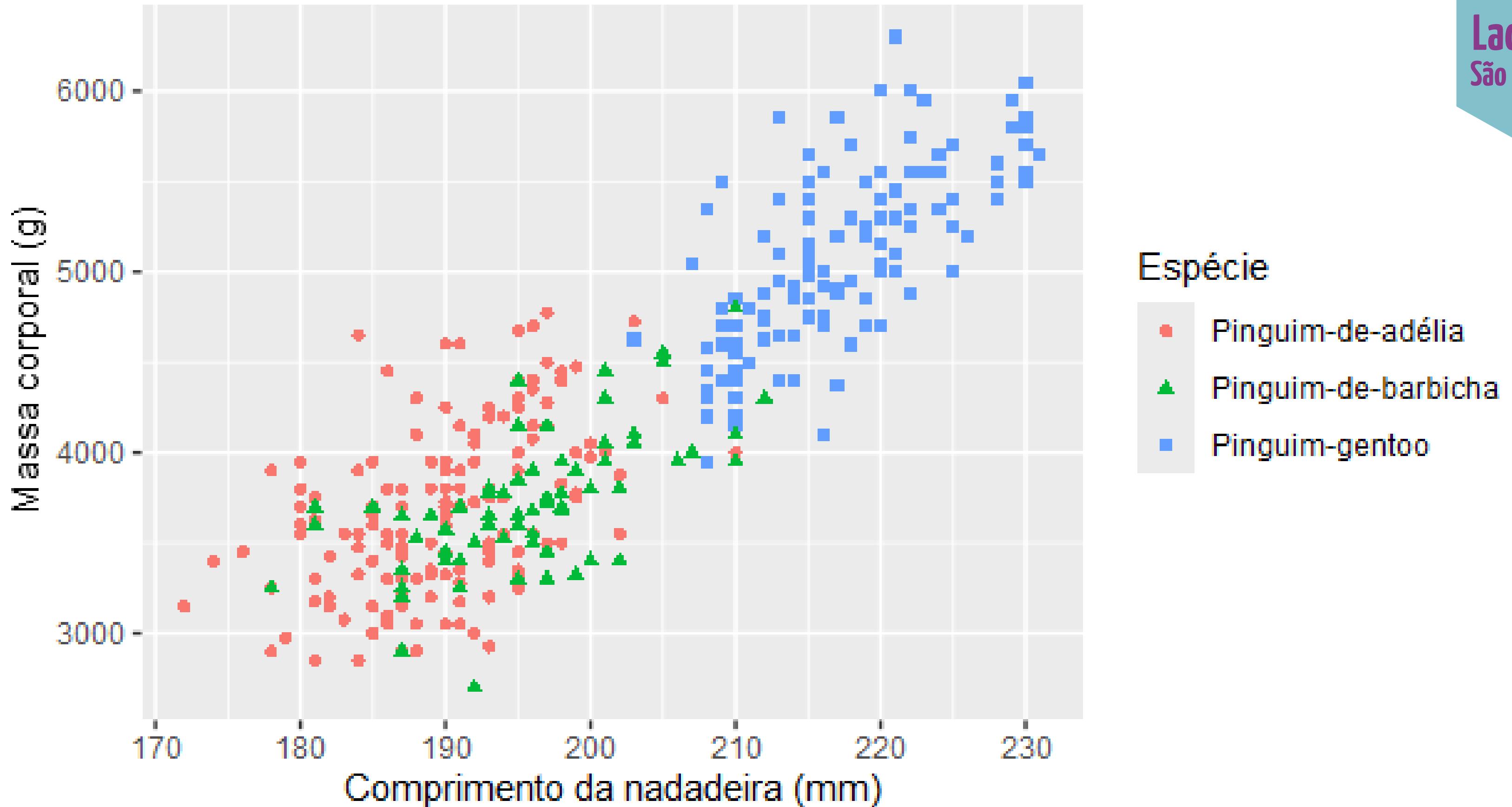
Categórica N macho, fêmea, fêmea, NA, fêmea, macho, fêmea, ma...

\$ ano

N. Discreta 2007, 2007, 2007, 2007, 2007, 2007, 2007, 2007, ...

Massa corporal e comprimento da nadadeira

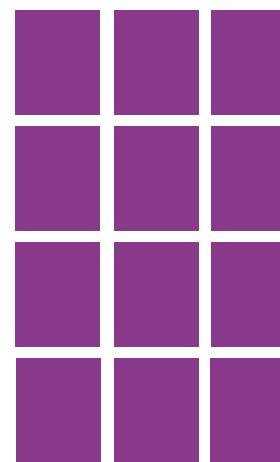
Medidas para Pinguim-de-adélia, Pinguim-de-barbicha e Pinguim-gentoo



Visualizando Dados Com GGPLOT

DADOS

`data`



STATistical
(transformações
estatísticas)



PROPRIEDADES ESTÉTICAS

`aesthetic`

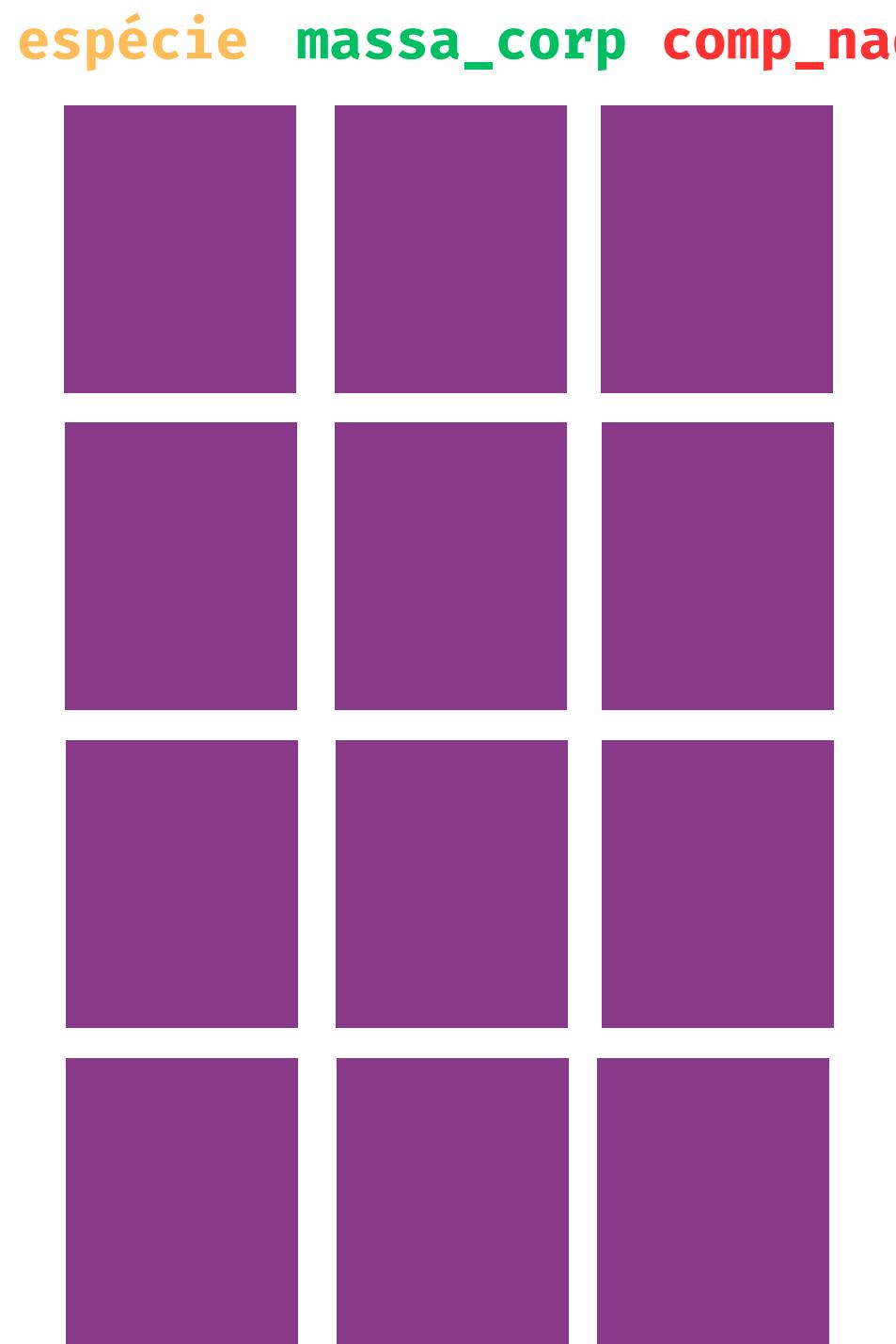
`x`
`y`
`color`
`fill`



GEOMETRIA

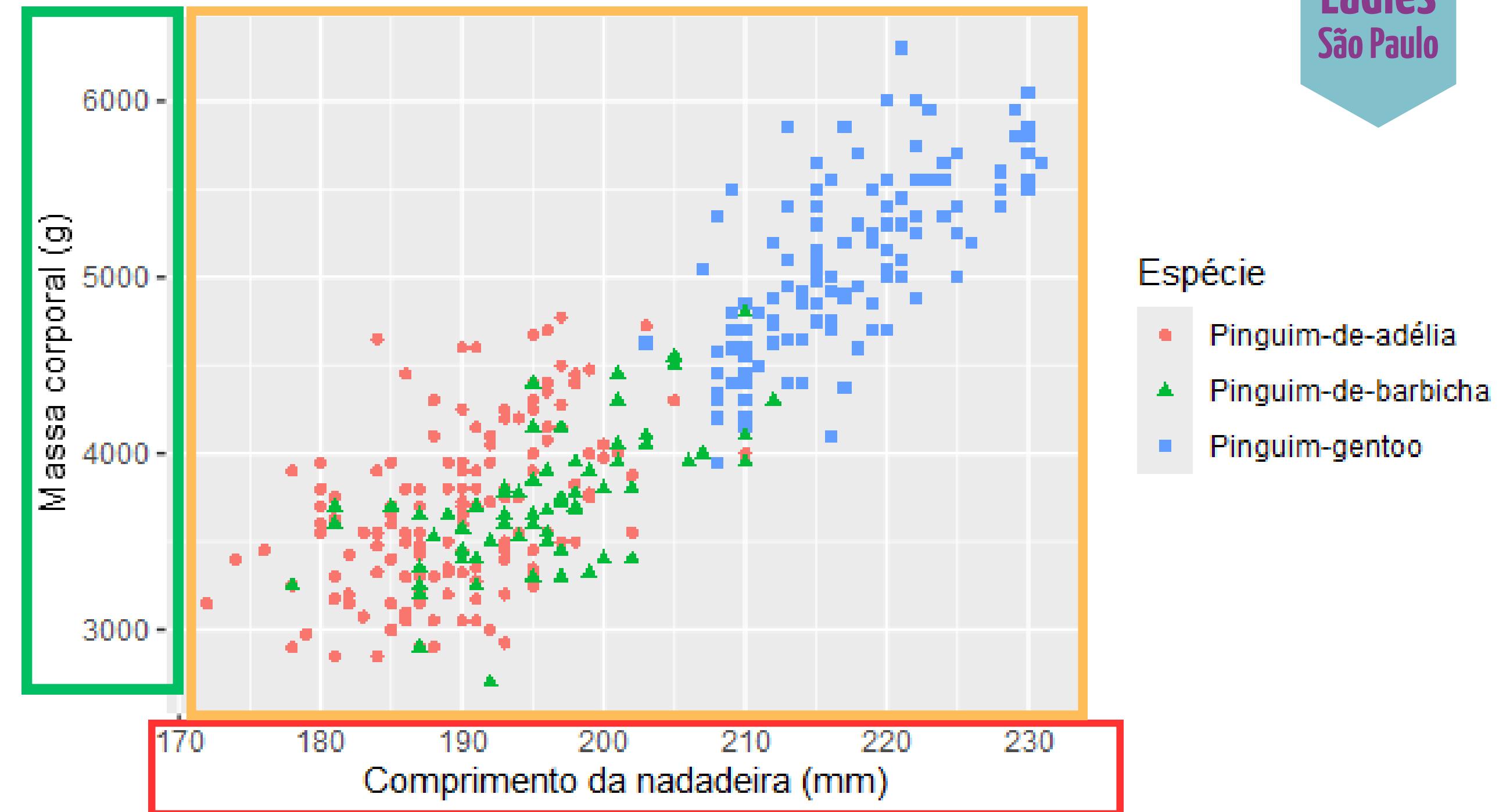
`geometries`

`geom_point()`
`geom_histogram()`
`geom_bar()`
`geom_col()`
`geom_line()`



Massa corporal e comprimento da nadadeira

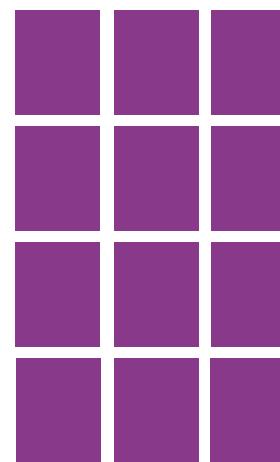
Medidas para Pinguim-de-adélia, Pinguim-de-barbicha e Pinguim-gentoo



Visualizando Dados Com GGPlot

DADOS

`data`



STATistical
(transformações
estatísticas)



PROPRIEDADES ESTÉTICAS

`aesthetic`

`x`
`y`
`color`
`fill`



GEOMETRIA

`geometries`

`geom_point()`
`geom_histogram()`
`geom_bar()`
`geom_col()`
`geom_line()`

Visualizando Dados Com GGPLOT

```
ggplot(  
  data = pinguins
```

Visualizando Dados Com GGPLOT

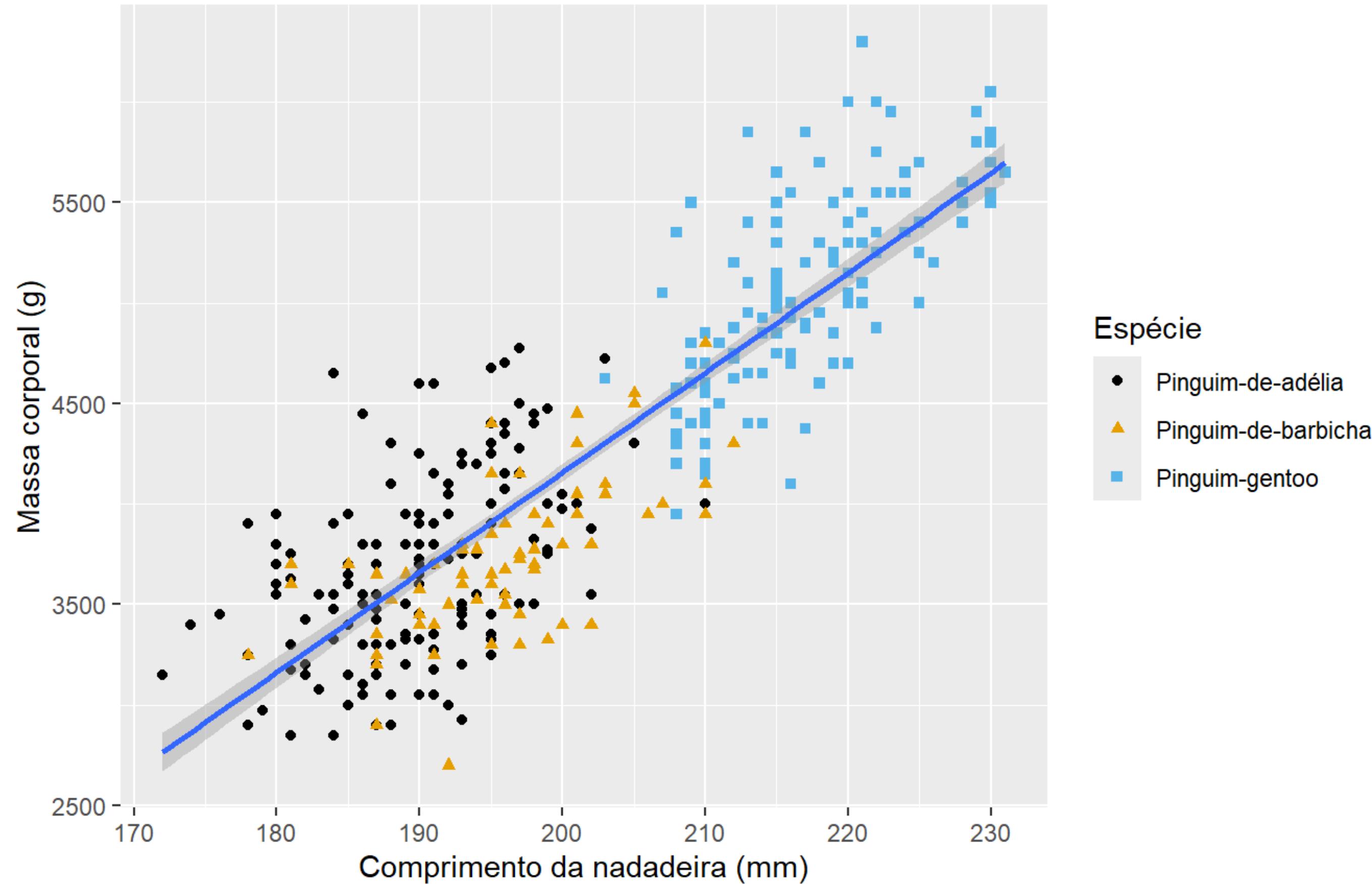
```
ggplot(  
  data = pinguins,  
  mapping = aes(x = comprimento_nadadeira,  
                y = massa_corporal))
```

Visualizando Dados Com GGPLOT

```
ggplot(  
  data = pinguins,  
  mapping = aes(x = comprimento_nadadeira,  
                 y = massa_corporal)) +  
  geom_point()
```

Massa corporal e comprimento da nadadeira

Medidas para Pinguim-de-adélia, Piguim-de-barbicha e Pinguim-gentoo



Exercícios

3. Faça um gráfico de dispersão de `profundidade_bico` em função de `comprimento_bico`. Ou seja, faça um gráfico de dispersão com `profundidade_bico` no eixo y e `comprimento_bico` no eixo x. Descreva a relação entre essas duas variáveis.

Exemplo

1. Ggplot

```
ggplot(data = penguins)
```

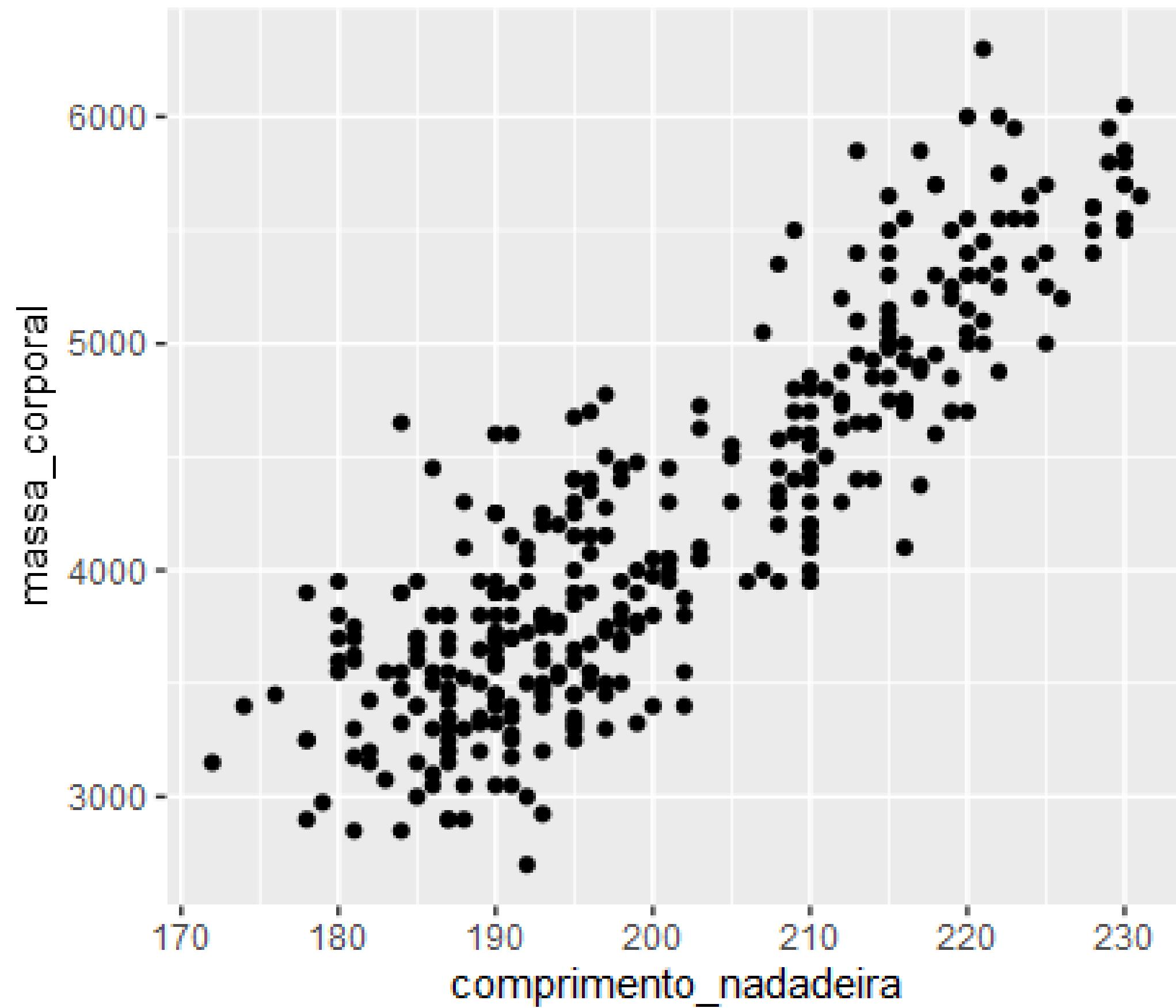
Exemplo

1. Ggplot

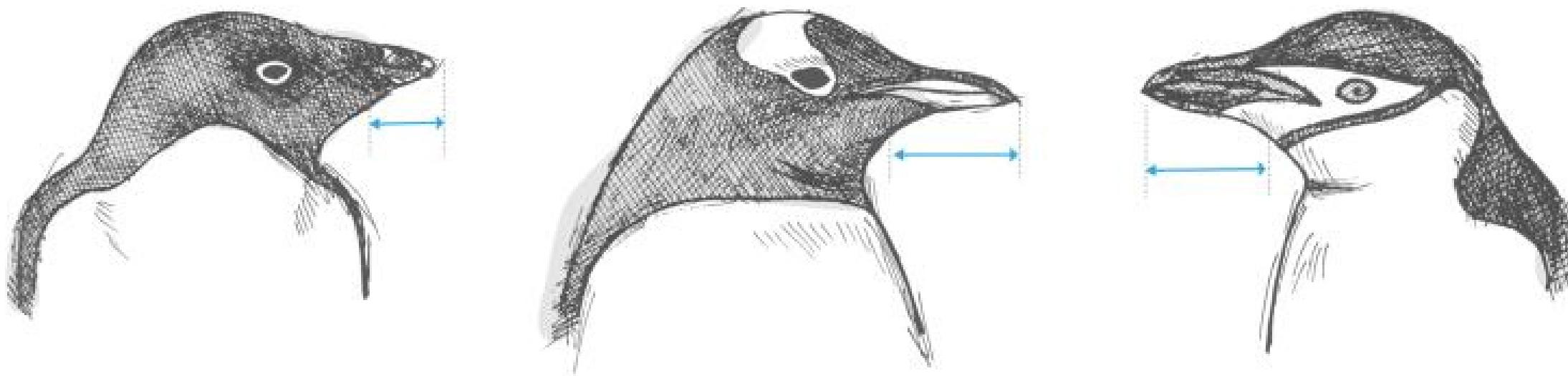
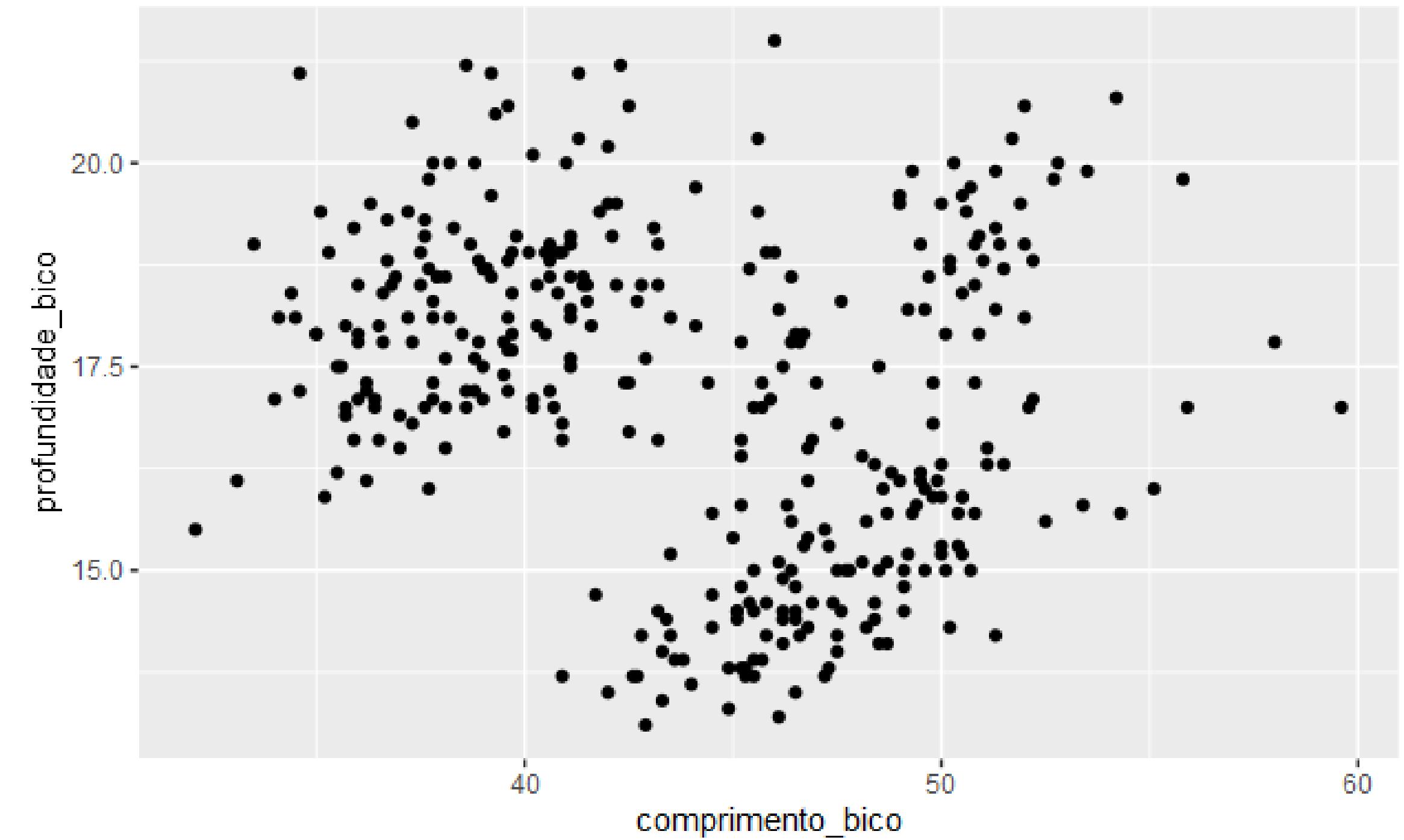
```
ggplot(data = pinguins)
```

2. Ggplot com data e com o mapeamento definindo as propriedades estéticas do seu gráfico

```
ggplot(data = pinguins,  
       mapping = aes(x = comprimento_nadadeira, y = massa_corporal) ) +  
       geom_point()
```



Exercício 3



Exercício

5. Por que o seguinte erro ocorre e como você poderia corrigi-lo?

```
ggplot(data = penguins) +  
  geom_point()
```

Exercício

5. Por que o seguinte erro ocorre e como você poderia corrigi-lo?

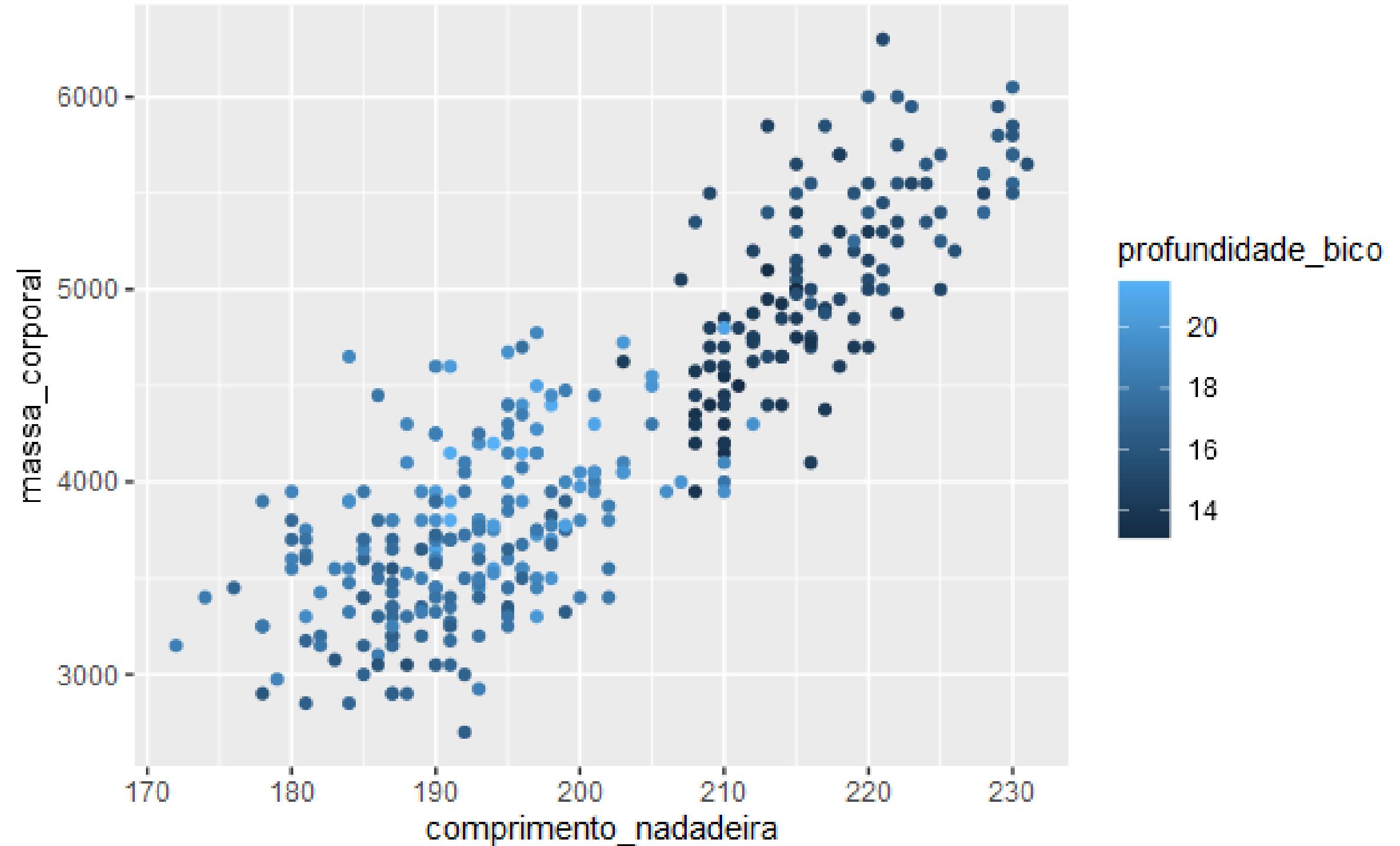
```
ggplot(data = pinguins) +  
  geom_point()
```

```
ggplot(data = pinguins,  
       aes(x = comprimento_bico, y = profundidade_bico)) +  
  geom_point()
```



Exercício (DESAFIO)

8. Recrie a visualização a seguir. Para qual atributo estético **profundidade_bico** deve ser mapeada? E ela deve ser mapeada na camada global ou na geometria?



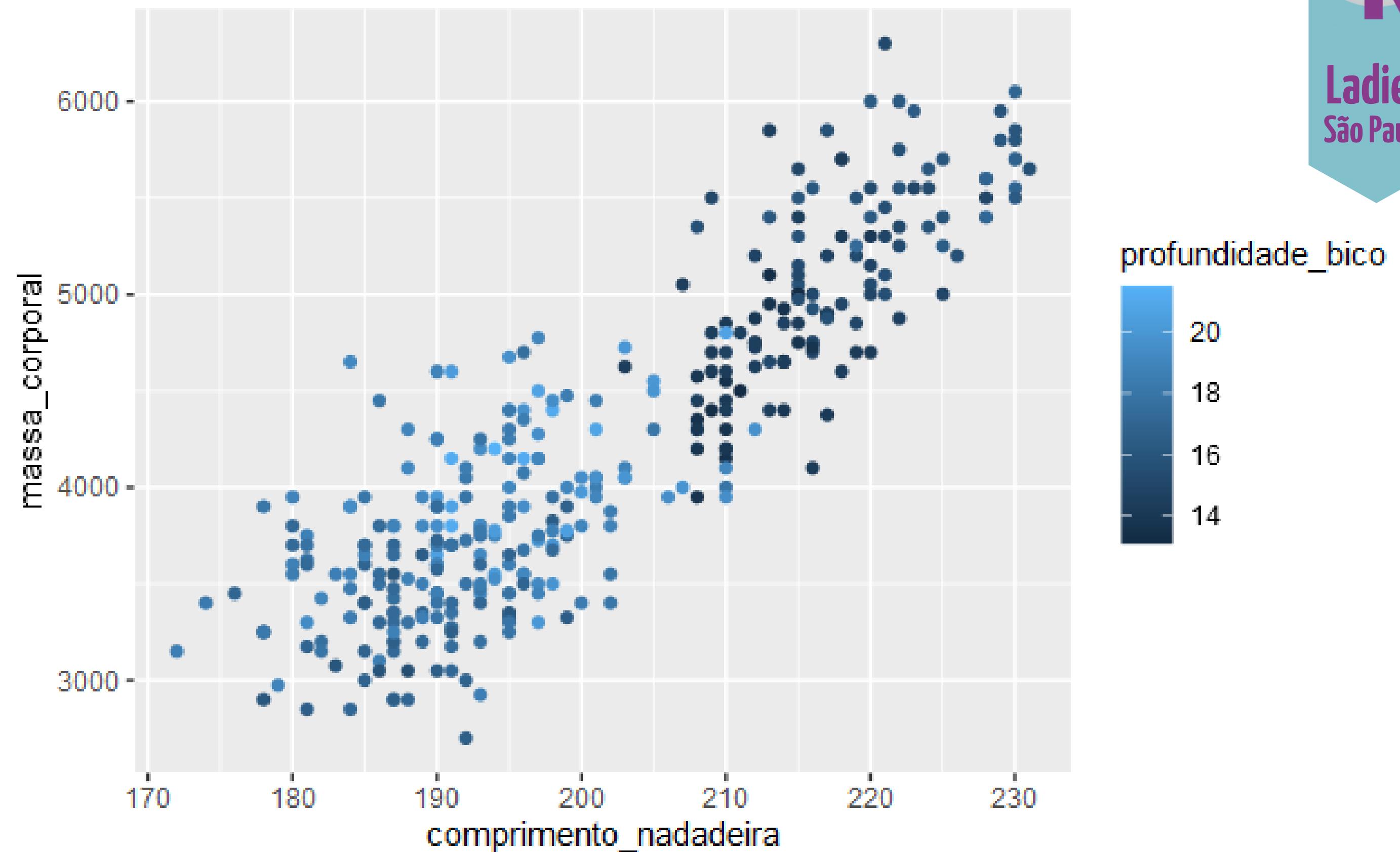
Exercício

8. Recrie a visualização a seguir. Para qual atributo estético profundidade_bico deve ser mapeada? E ela deve ser mapeada no nível global ou no nível da geometria?

```
ggplot(pinguins, aes(comprimento_nadadeira, massa_corporal)) +  
  geom_point(aes(color = profundidade_bico))+  
  labs(color = "profundidade_bico")
```

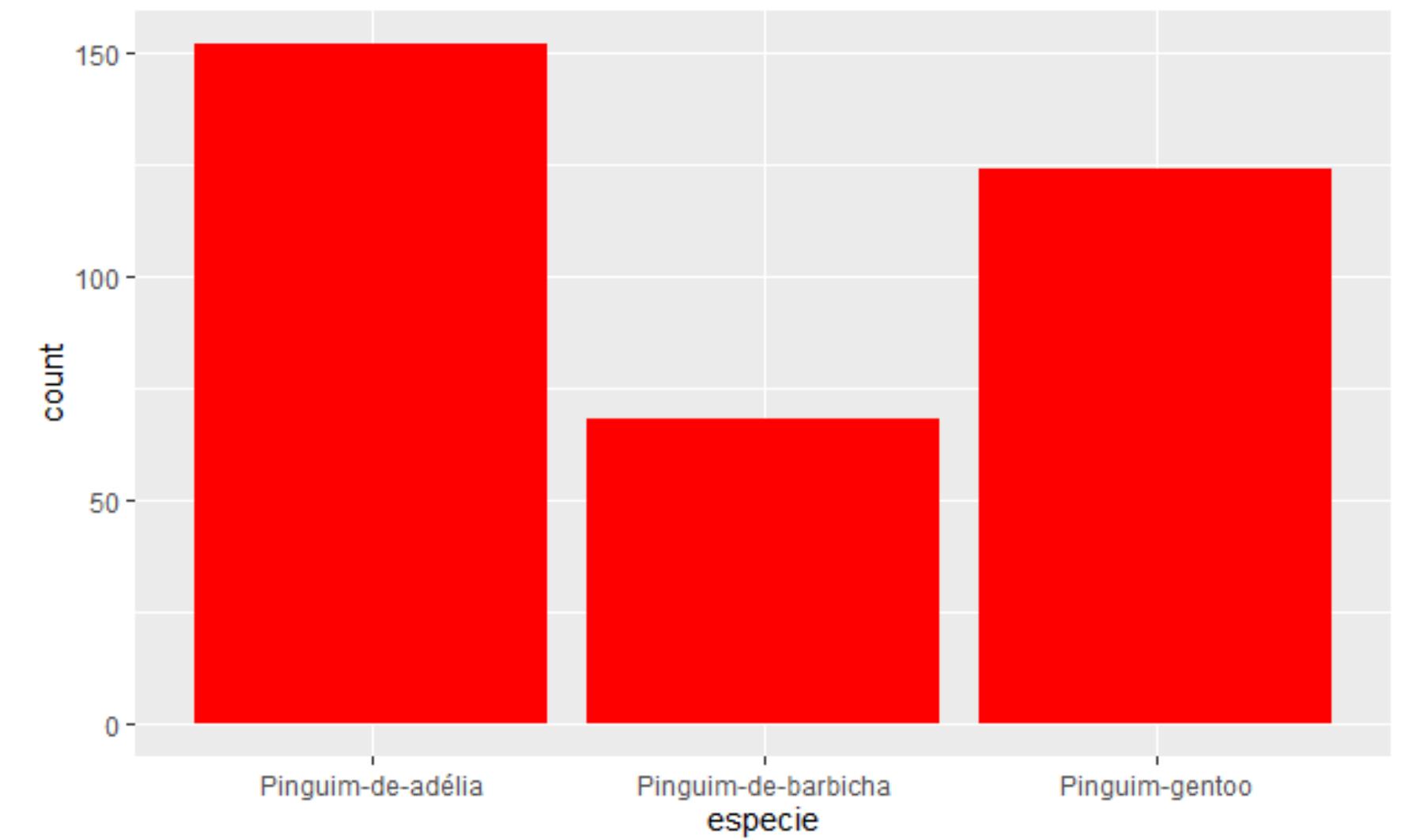
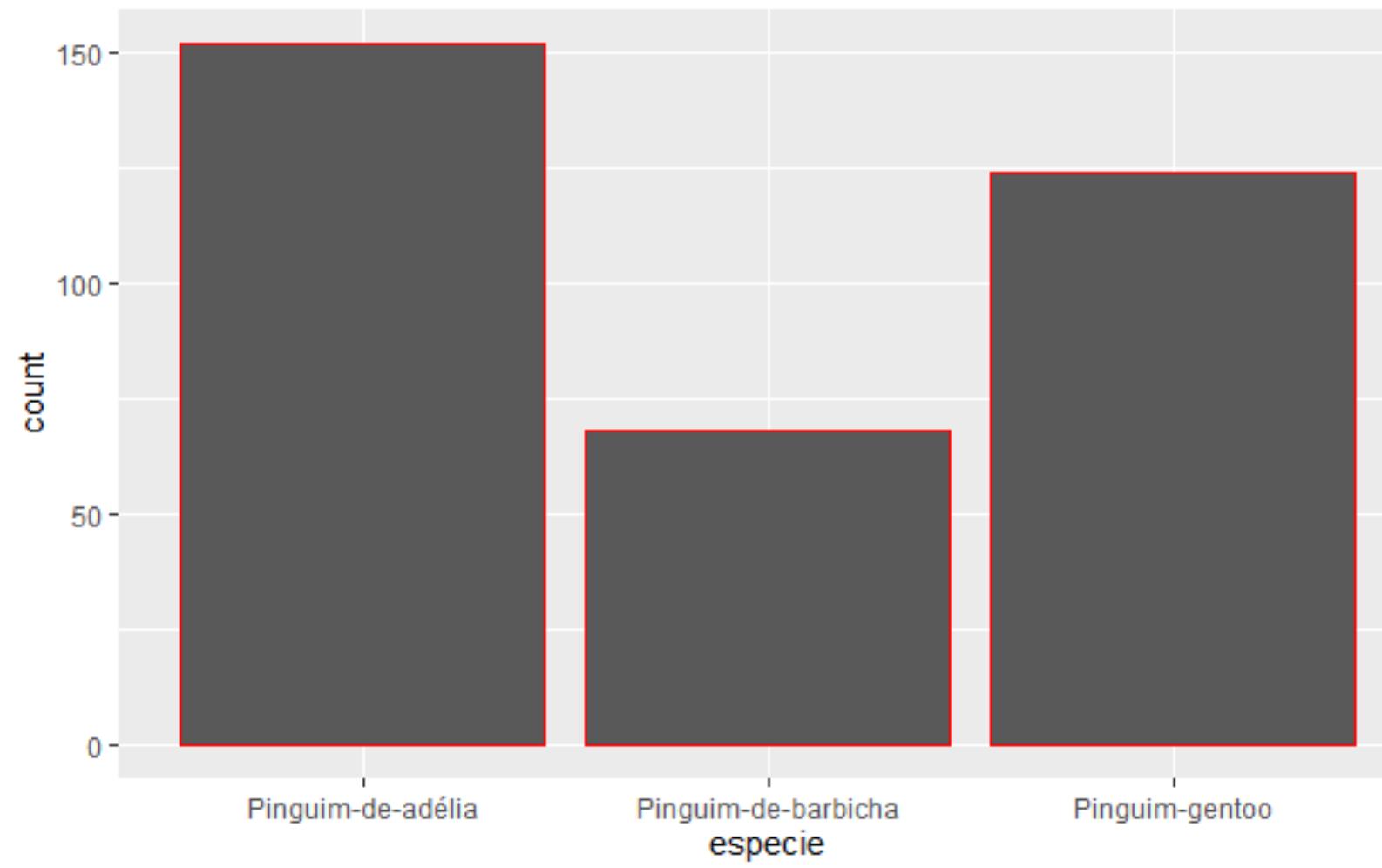


Exercício



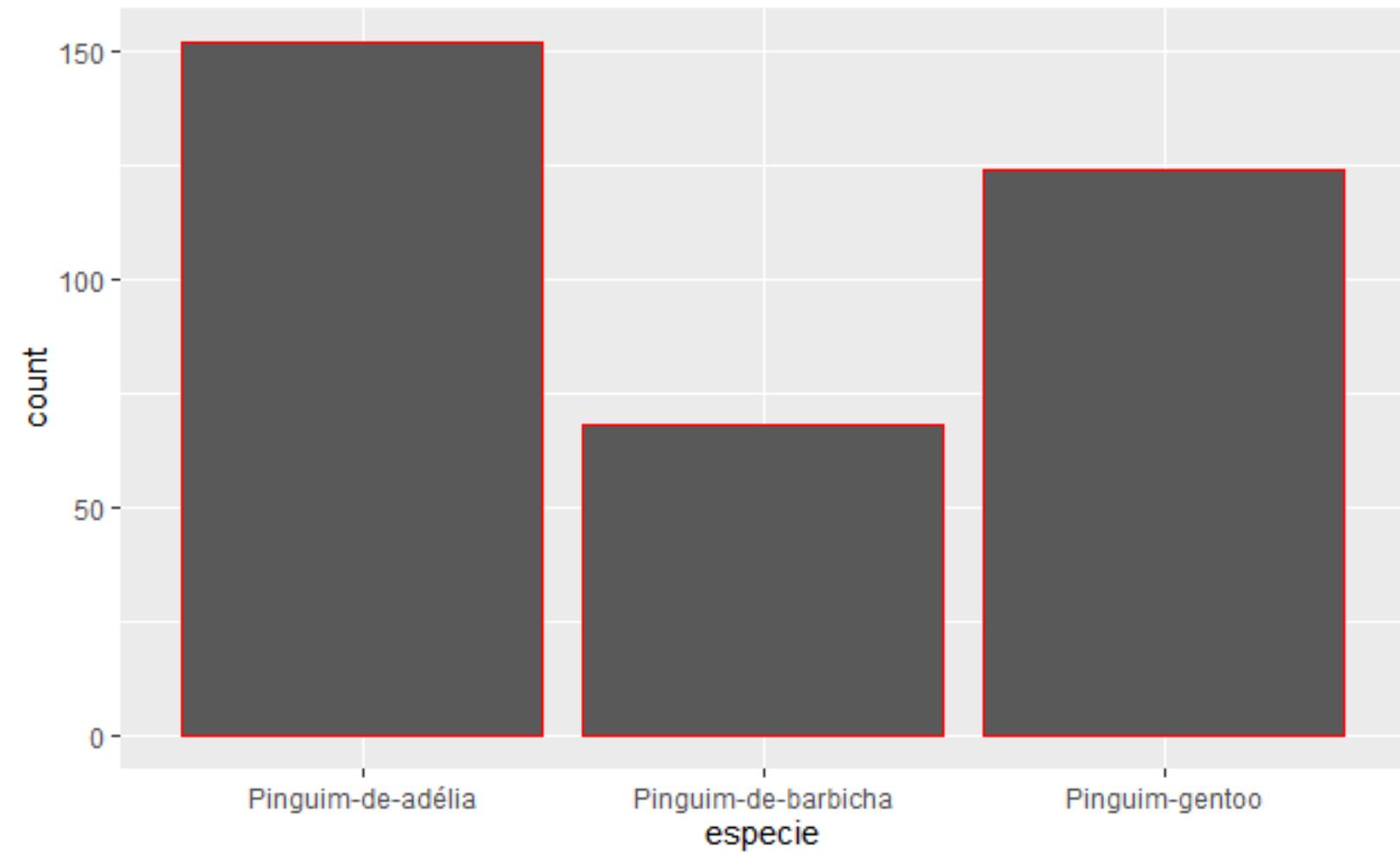
Exercício 2

2. (1.4.3) Como os dois gráficos a seguir são diferentes? Qual atributo estético, color ou fill, é mais útil para alterar a cor das barras?



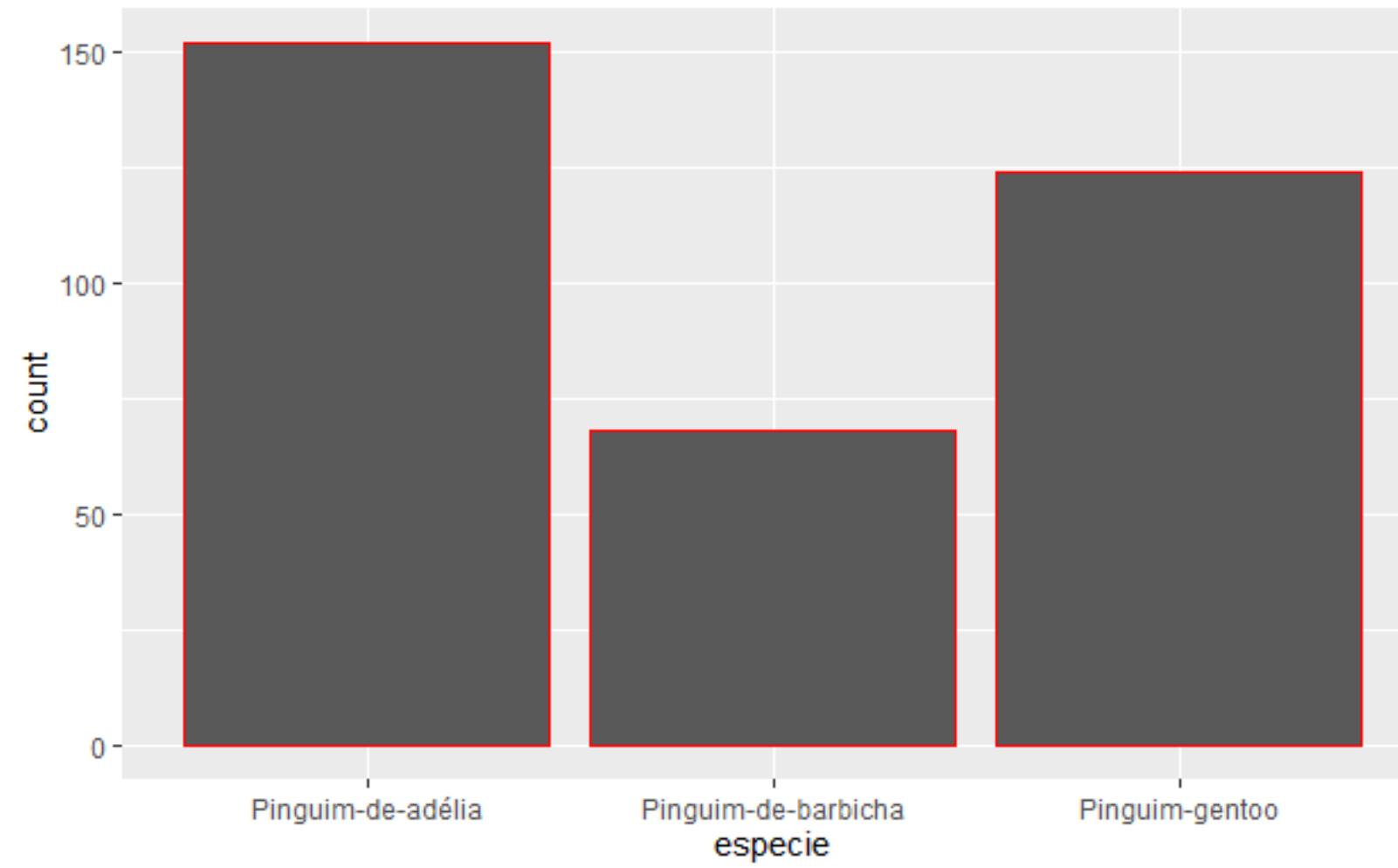
Exercício 2 (1.4.3)

```
pinguins|>  
ggplot(aes(x= especie))+  
geom_bar(color = "red")
```

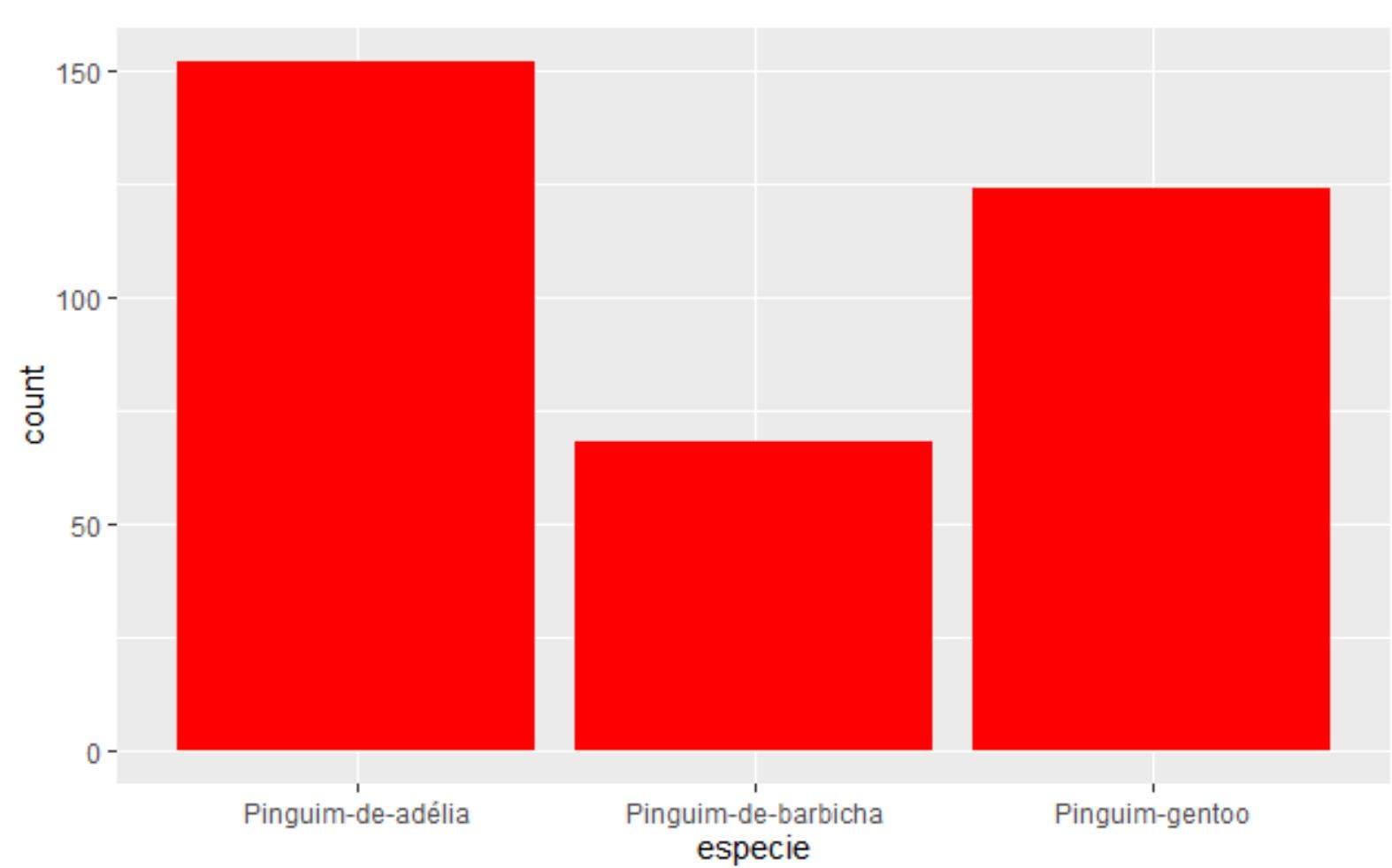


Exercício 2 (1.4.3)

```
pinguins|>  
ggplot(aes(x= especie))+  
geom_bar(color = "red")
```

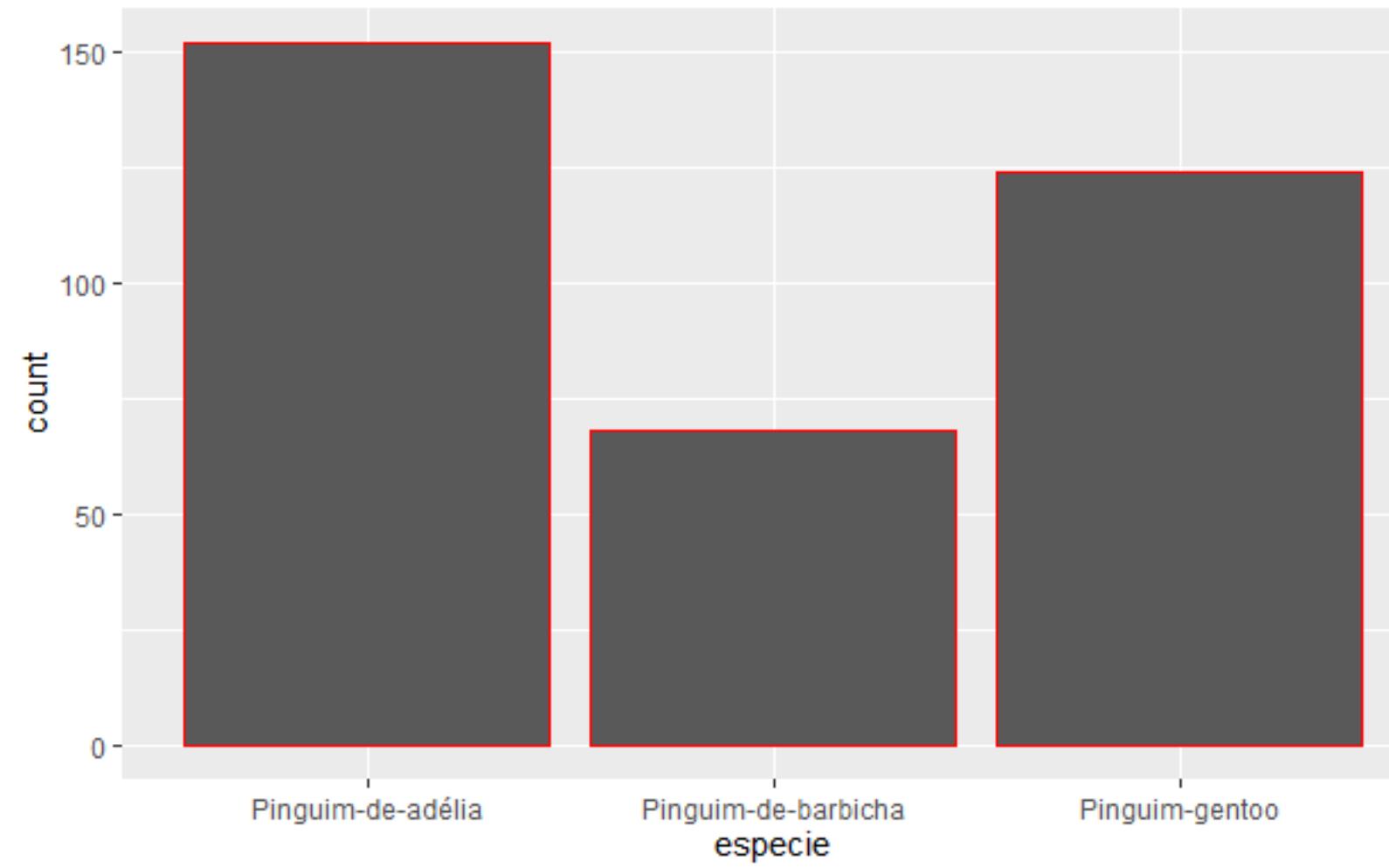


```
pinguins|>  
ggplot(aes(x= especie))+  
geom_bar(fill="red")
```

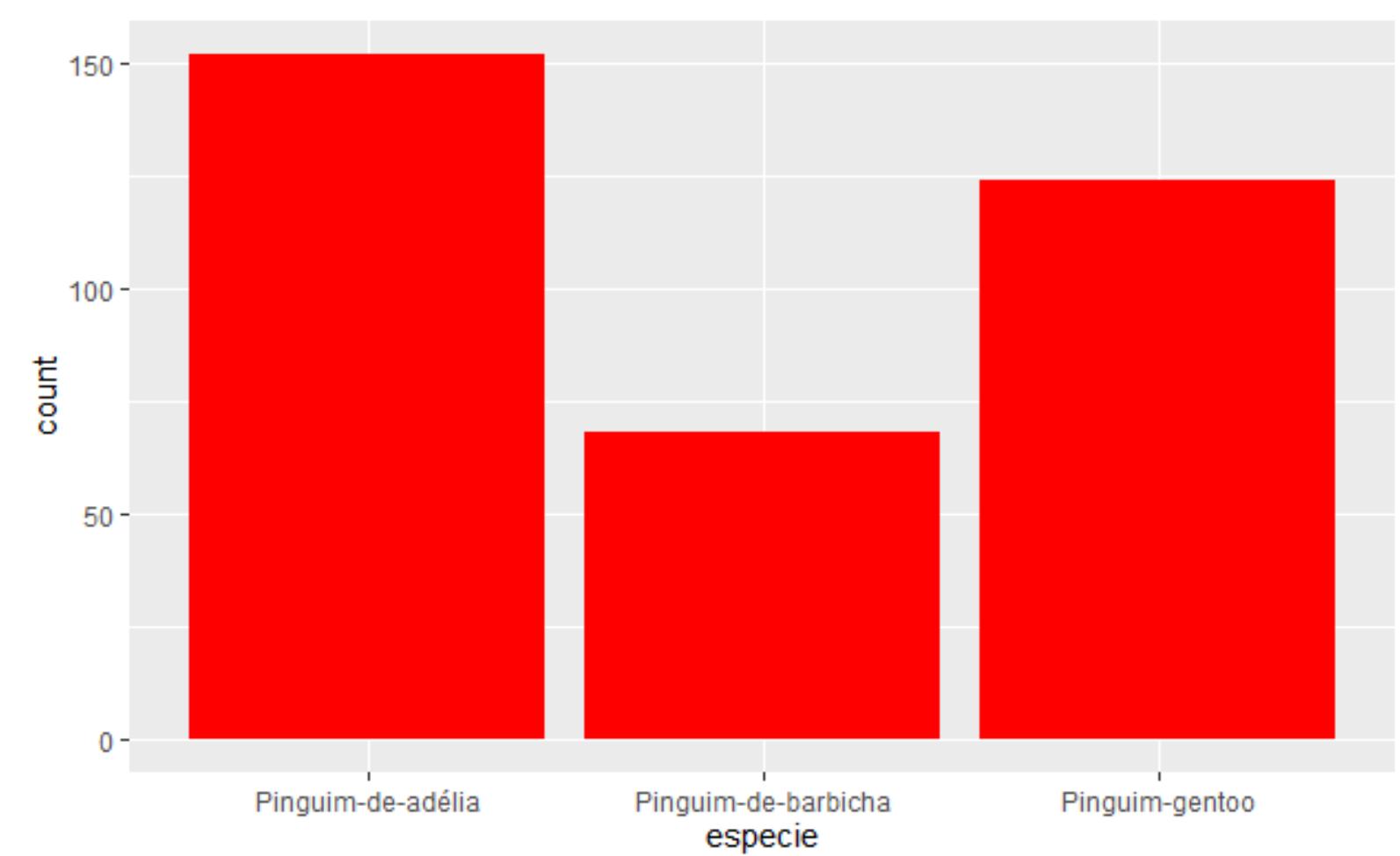


E se mudar color = especie ?
E se mudar fill = especie ?

```
pinguins|>  
  ggplot(aes(x= especie))+  
    geom_bar(color = "_____")
```

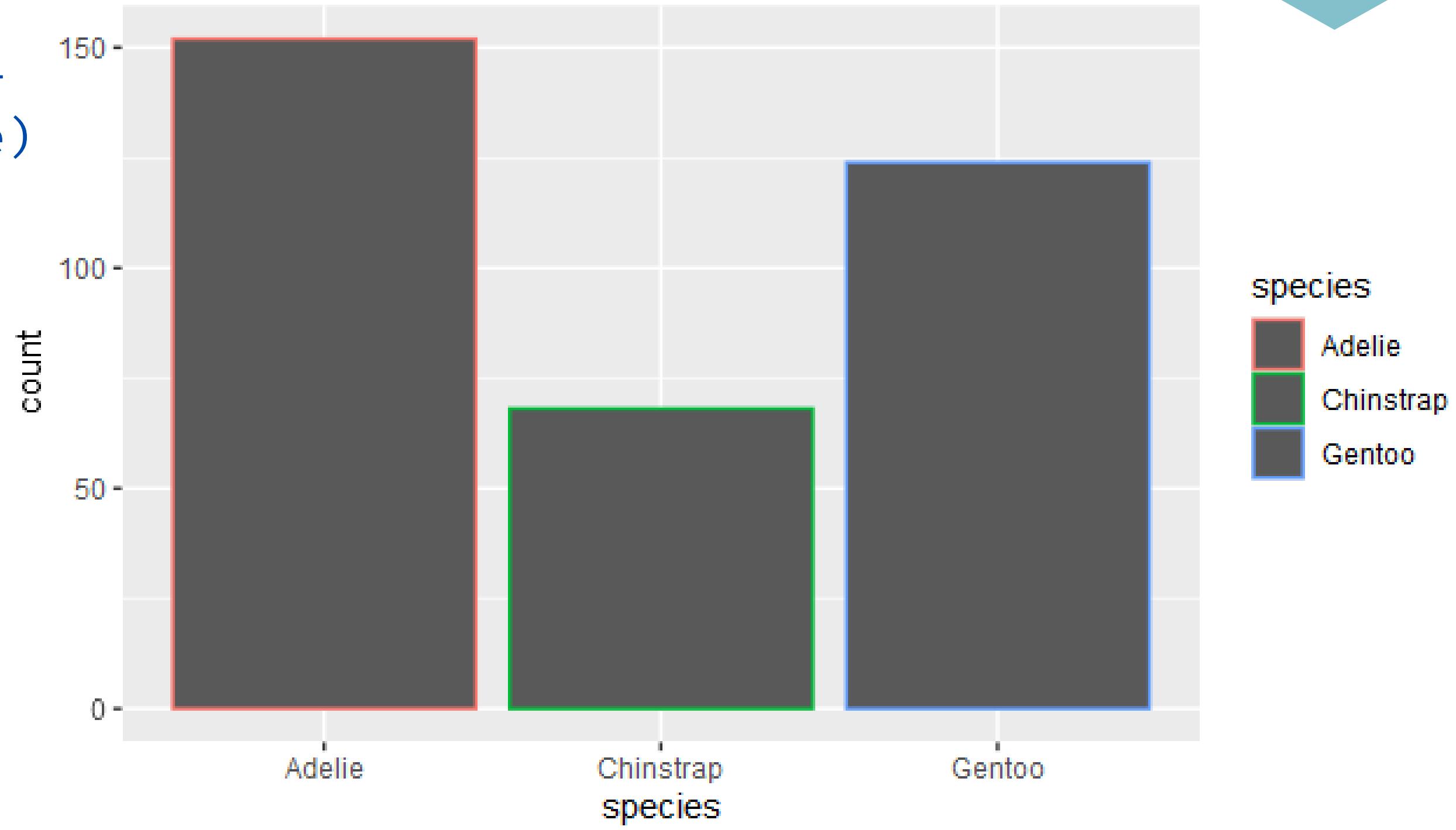
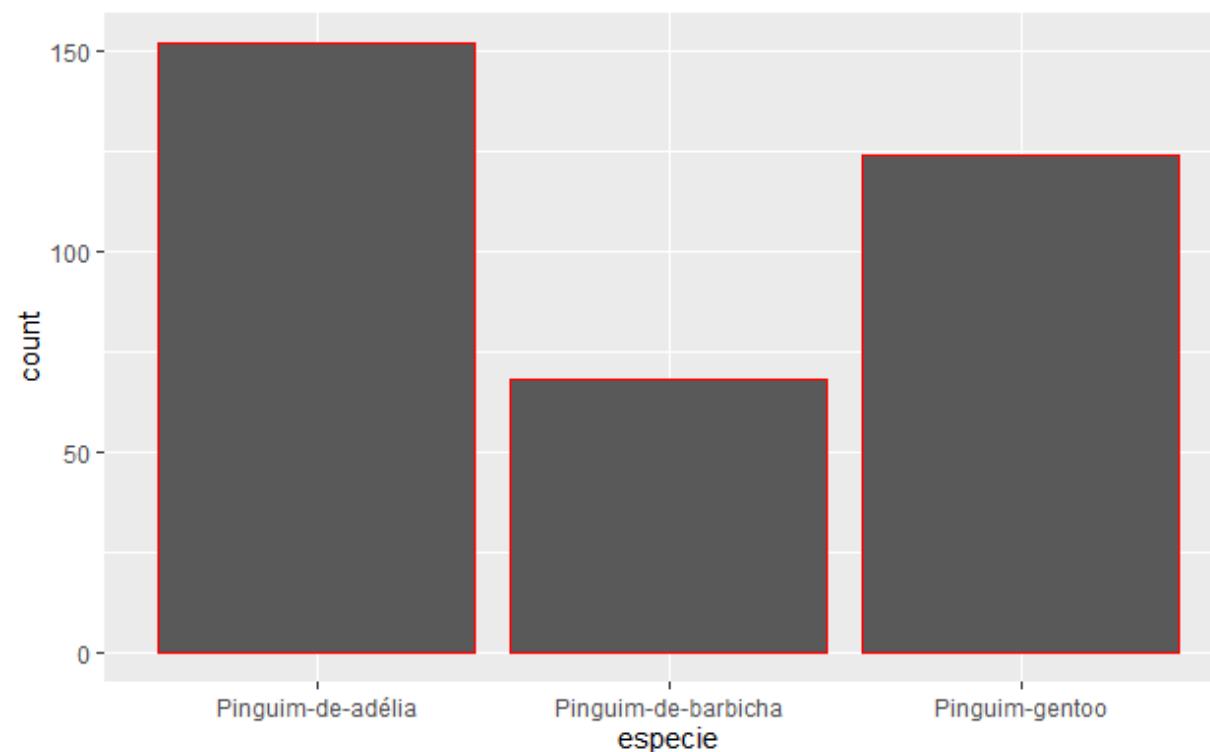


```
pinguins|>  
  ggplot(aes(x= especie))+  
    geom_bar(fill="_____")
```



E se mudar color = especie ?
E se mudar fill = especie ?

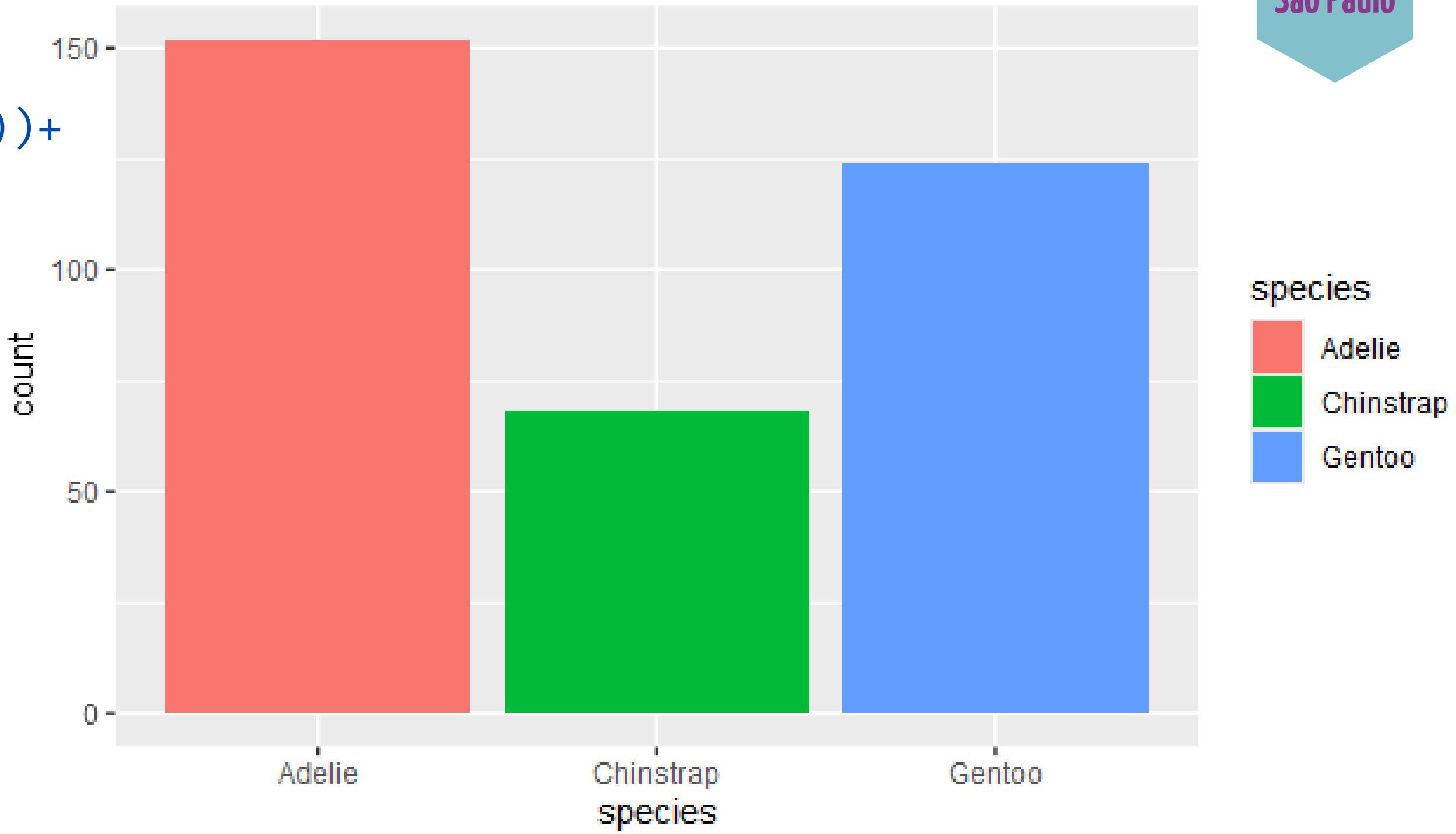
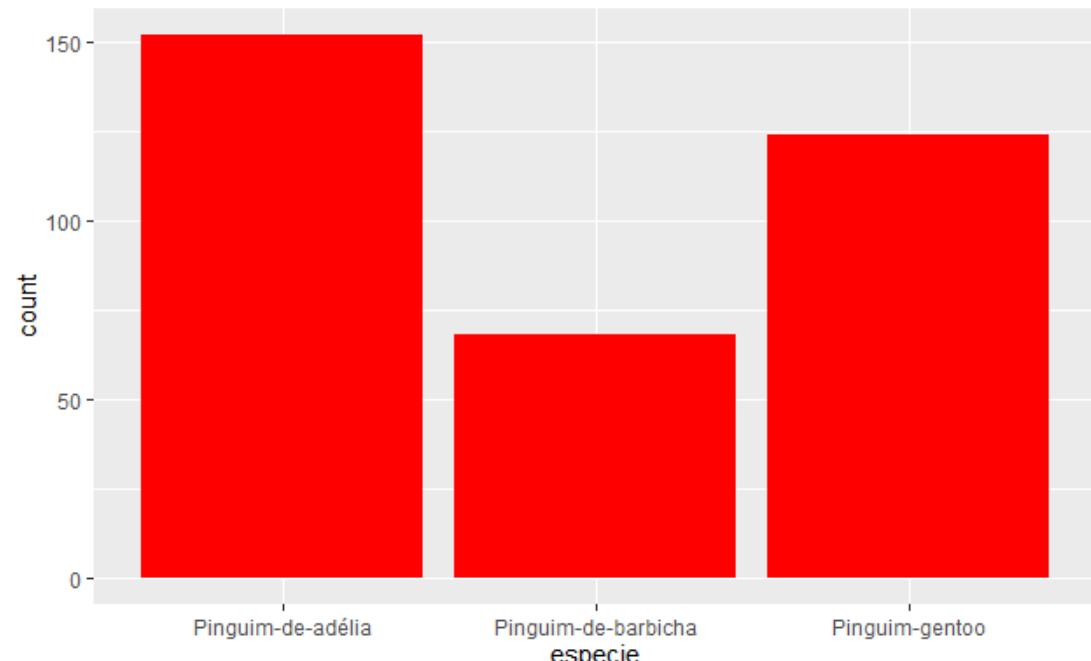
```
pinguins|>
ggplot(aes(x= especie))+
geom_bar(color = especie)
```



E se mudar fill = especie ?

```
pinguins|>
  ggplot(aes(x= especie))+  

    geom_bar(fill="_____")
```

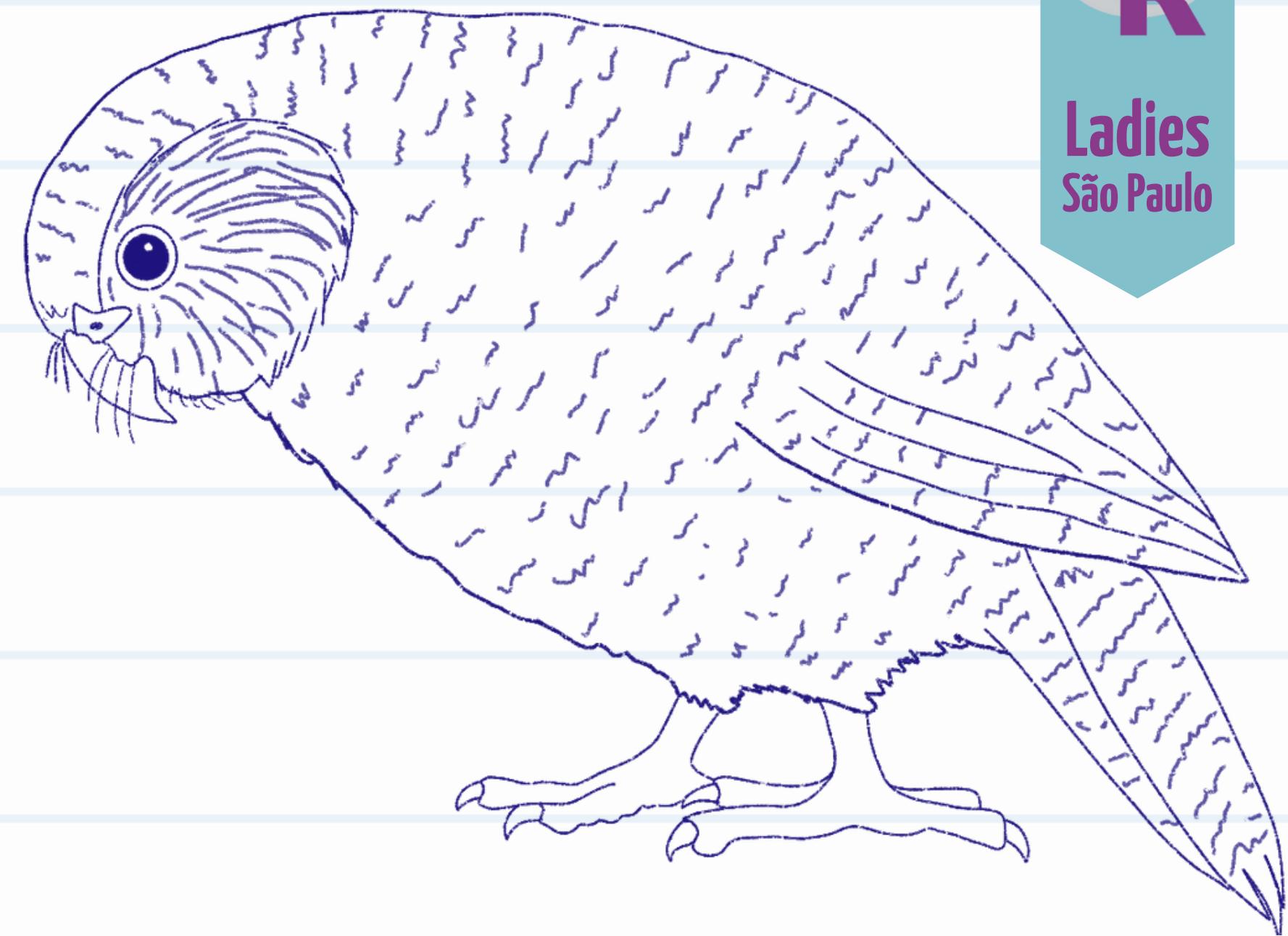


Dicas Para Problemas Comuns

- Compare seu código com o código do livro;
- Observe se todos parênteses estão completos;
- Se ver o + no seu console, o R está aguardando uma expressão completa;
- O posicionamento do + no ggplot no final da linha;
- Use o painel Help para obter ajuda ou `?nome_da_função` no console;
- Leia atentamente o erro, e faça uma busca no google.

Resumo

- uma visualização é uma estrutura em camadas;
- há vários tipos de atributos estéticos como posição (position), cor (color), tamanho (size) e forma (shape);
- podemos aumentar a complexidade e melhorar a apresentação dos gráficos camada por camada;
- todos esses detalhes são aprofundados a medida que o livro avança.



Dúvidas?

Próximo encontro:

Facilitação: Fernanda Peres

Capítulos:

- Fluxo de Trabalho: básico
- Fluxo de trabalho: scripts e projetos



Data: 22/04/2024